

Multicriteria Ranking Using Weights Which Minimize the Score Range

Chris Tofallis

Abstract Various schemes have been proposed for generating a set of non-subjective weights when aggregating multiple criteria for the purposes of ranking or selecting alternatives. The maximin approach chooses the weights which maximise the lowest score (assuming there is an upper bound to scores). This is equivalent to finding the weights which minimize the maximum deviation, or range, between the worst and best scores (minimax). At first glance this seems to be an equitable way of apportioning weight, and the Rawlsian theory of justice has been cited in support.

We draw a distinction between using the maximin rule for the purpose of assessing performance, and using it for allocating resources amongst the alternatives. We show that it has a number of drawbacks which make it inappropriate for the assessment of performance. Specifically, it is tantamount to allowing the worst performers to decide the worth of the criteria so as to maximise their overall score. Furthermore, when making a selection from a list of alternatives, the final choice is highly sensitive to the removal or inclusion of alternatives whose performance is so poor that they are clearly irrelevant to the choice at hand.

1 Introduction

One of the most influential works in the area of moral and political philosophy in the last 50 years has been John Rawls's *A Theory of Justice* (1971). Rawls rejects the utilitarian idea of "the greatest good for the greatest number". This is a concept which the multi-criteria decision community would recognize as being fraught with difficulties. These include the fact that "the good" is likely to be a multi-factor concept, and that we are also dealing with multiple stakeholders holding different views. It is important to note that even if there were agreement on how to measure

C. Tofallis

The Business School, University of Hertfordshire, College Lane, Hatfield,
Hertfordshire AL10 9AB, UK
e-mail: c.tofallis@herts.ac.uk

and then aggregate the overall good of the population, it does not follow that maximizing it would provide any form of social justice unless of course such justice was built into the definition of “the good”. Rawls viewed “justice as fairness” and felt that the worst off should not be made even worse. In particular, if public resources are to be distributed unequally, then the worst off should benefit the most. Rawls referred to this as the “difference principle”.

Rawls has been cited in support of using the maximin rule for weighting criteria by Pettypool and Karathanos (2004). They proposed the rule for the purpose of appraising the work of employees under a number of criteria. Butler and Williams (2002) use the maximin rule in sharing out the fixed costs associated with shared facilities. In support of it they cite work based on experiment and survey:

A variety of fairness criteria are discussed in the seminal paper of Yaari and Bar-Hillel (1984). They conducted a series of experiments to see which of nine possible criteria were considered most fair by a sample of people questioned. In relation to needs, an allocation based on minimizing the maximum inequality was overwhelmingly considered the most fair.

One field where the minimax concept is widely used is in location problems. When choosing locations for emergency facilities (police, ambulance, firefighting) or other public offices or services, this method selects locations so as to minimize the maximum travel time or distance to any person who is being served. The method has been criticized (e.g. Ogryczak 1997) because if there is a single recipient (or a small cluster) that is located far from the vast majority, then a location may be selected which is far from all recipients. There is thus seen to be a disproportionate effect on the decision by a tiny minority of the recipients. We shall see that a similar difficulty arises when applying the minimax concept to multicriteria weighting.

The minimax objective is also used as an alternative to least squares in regression. It involves minimizing the largest deviation or residual. This is an appropriate objective if the error distribution is uniform; this can arise when the errors arise as a result of rounding, e.g. a digital measurement device will have a limited number of digits to display. This type of regression is not appropriate if there are outliers in the data, as these will severely distort the resulting model.

Another application of the maximin objective is in the allocation of highway patrol officers to districts so as to ensure that all districts experience a reduction in speeding; the aim was to maximize the minimum reduction in the number of speeding offences (Rardin 1998, p. 158). In the field of scheduling jobs numerous objectives are used, one of these is to minimize the maximum lateness (Rardin 1998, p. 605). It is also used to minimize maximum congestion or bottlenecks. Du (1996) surveys the field of minimax applications.

2 Geometric Representation

The maximin concept has been used in the assessment of performance by a number of authors. For example, Karsak and Ahiska (2005) and Karsak (2004) consider the problem of attaching weights to the various outputs (criteria of the type “more is better”) when there is a single input. To create an efficiency score each output is

divided by the input and then weights are attached to each of these ratios. In DEA (data envelopment analysis) each alternative has its own weights. These are chosen so as to optimize the score for that alternative. Because this method attaches different weights for each alternative, this leads to the generation of an efficient frontier which is made up of piecewise linear segments. In DEA all the alternatives on the frontier are given the same score of 100%. In an effort to increase the discrimination between such units and identify a preferred alternative, they seek a common set of weights to be used across all alternatives. These non-negative weights are chosen so as to maximize the minimum score (maximin), subject to the condition that all scores do not exceed 100%. In criterion space a set of common weights corresponds to a line or plane.

Figure 1 shows an example involving two criteria. According to DEA, points A, B and C are ranked first with the maximum score, and ABC delineates the DEA frontier. In DEA alternative P has a score given by the ratio OP/OP' , where P' is the point where the ray OP intersects the frontier. Because P' lies between A and B, the corresponding weights are determined by the slope of the line AB. Point T however would be assessed relative to the line segment BC, which corresponds to a different set of criteria weights. Of the points shown in Fig. 1, P would have the lowest score. If we now depart from the piecewise frontier in favour of a single set of common weights based on the maximin rule, we shall have a single extended line frontier. We shall have to choose weights which maximize P's score, and so the frontier will be AB (extended). Notice that the particular line segment and hence weights, are chosen by reference to the worst performing alternative. This in itself is strange because the frontier is supposed to represent best practice, and yet its location is crucially influenced by an alternative displaying worst practice.

Troutt et al. (1993) use the maximin rule as a way of further ranking those alternatives which have all been given the same 100% efficiency score from a data envelopment analysis. This differs from the above in that only efficient alternatives are considered at this second stage. Hence the worst performers cannot influence

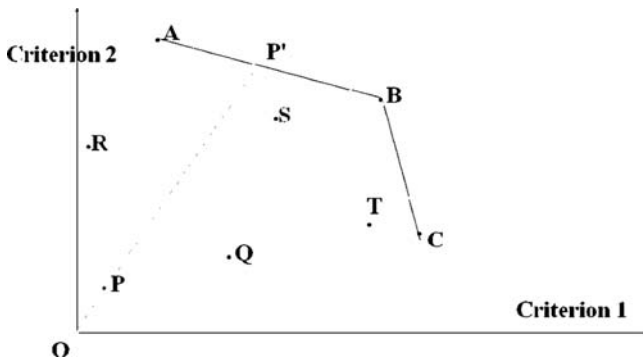


Fig. 1 Having a common set of weights (with an upper limit to the overall score) means that a line such as AB or BC acts as the frontier. The slope of such a line determines the weights

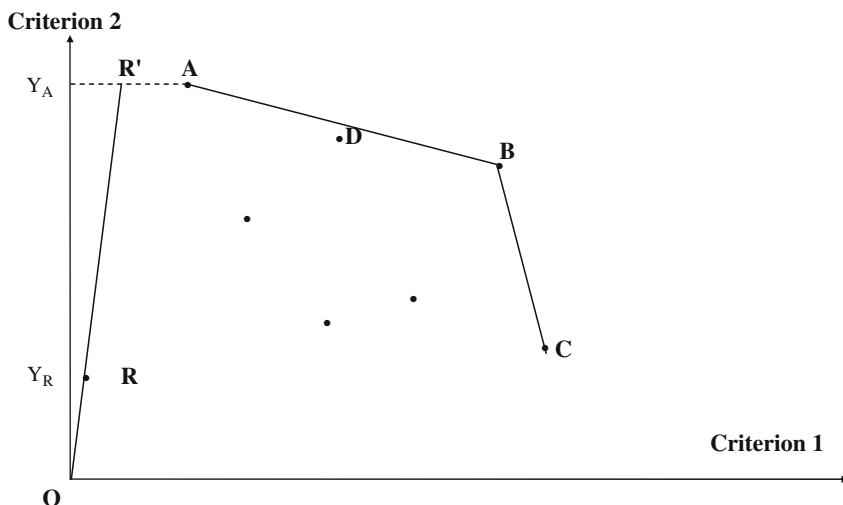


Fig. 2 When we attempt to project alternative R onto the frontier we find that its “target” (R') does not lie between observed efficient units – i.e. it is not naturally enveloped. This leads to a horizontal frontier and a zero weight for criterion 1

the resulting weights. This is a definite improvement. The alternatives which will now influence the position of the linear frontier will be those that are at the ends of the frontier. In a two dimensional setting these will be points A and C, but in higher dimensions they will be the points on the perimeter of the frontier. Such points have very high scores in one criterion but are weak in the others, and are sometimes referred to as “mavericks”. They contrast with good all-rounders. One might also include in this second stage those alternatives which are Pareto-optimal even though they do not appear on the convex hull, for example point D in Fig. 2. Such points are also “good all-rounders”.

Now consider what happens when alternative P is removed from Fig. 1. Q now has the lowest score. This forces facet BC (extended) to act as the new frontier. Unit A was previously ranked first equal (maximum score), but now it slides down the rankings below B, C, T, S and Q! Karsak and Ahiska (2005) used the maximin method in a selection problem: to choose a particular piece of equipment from a number of competing alternatives. Expressed in these terms the removal of a point such as P corresponds to removing an irrelevant alternative – one that would never be selected because of its poor performance. Yet its removal causes huge changes in the rankings. This violates the axiom of decision theory known as Sen’s property alpha (Sen 1969), also known as the Chernoff condition (Chernoff 1954), which states that the removal or addition of an irrelevant alternative should not affect the decision. The selection decision should be independent of irrelevant alternatives. The removal of such unwanted points could for example arise in an initial screening stage, where alternatives which do not measure up to certain minimum standards are removed from further consideration. They could also be removed from simple dominance

arguments. A memorable illustration of the principle is an anecdote attributed to the philosopher Sidney Morgenbesser:

After finishing dinner, Sidney Morgenbesser decides to order dessert. The waitress tells him he has two choices: apple pie and blueberry pie. Sidney orders the apple pie. After a few minutes the waitress returns and says that they also have cherry pie, at which point Morgenbesser says “In that case I’ll have the blueberry pie.”

Troutt (1997, and references therein) has written a number of papers applying the maximin approach to DEA with both multiple inputs and multiple outputs. He calls the resulting scores the MER – the maximin efficiency ratios. He makes the following observation:

When the MER model was first discussed (without subsequent benefit of theoretical justification), some critics argued that “optimal” multipliers should not be based on least efficient units. While that criticism has intuitive merit, it may be noted that a reverse perspective is actually more fruitful. Namely, the minimum efficiency, as well as the average (or any other summary statistic) depends on the weights. Such weights or multipliers may, or may not, in general, maximize the likelihood of the resulting aggregate measure. Thus, from the maximum likelihood perspective the procedure appears intuitive. However, this apparent “contradiction of intuitions” continues to be interesting and not yet fully resolved.

Troutt and Zhang (1993) also note that “a possible objection is that the resulting weights may be overly influenced by the worst performers”. They try to address this by saying “choices of weights which increase the minimum ratio frequently increase the average ratio as well, and conversely. Hence the maximin aggregation principle appears similar in expected performance to maximization of the average, which clearly depends on the performance data of the whole set of [alternatives]”. This is not a persuasive argument because in the maximization of the average each point has equal influence, whereas in the maximin case this is far from being true. They also try to address the issue by first noting that using maximin leads to all scores being squeezed into the narrowest range – which is true. It is then argued that the range is a measure of dispersion, as is the variance, so one would expect similar performance to minimizing the variance of the scores, and variance does depend on all of the data. Once again, this conclusion does not follow because the calculation of variance is based on all observations whereas the range is not.

To help us understand why we would not expect similar scoring performance let us draw some parallels with methods of fitting models to data. Consider the deviations from the 100% score as being residuals, and consider that we are fitting a linear model which is constrained not to have any data points lying above it. It now becomes clear that the maximin approach corresponds to fitting using the Chebyshev or L_∞ norm, and the minimization of the average residual corresponds to the L_1 norm. It is well established that these fitting approaches produce very different models and so we cannot expect to obtain similar performance as claimed above. Specifically, the L_1 norm is less sensitive to outliers than least squares regression, whereas the Chebyshev norm is more sensitive to outliers than least squares.

3 Can the Maximin Approach Produce a Single Winning Alternative?

Pettypool and Karathanos (2004) propose the maximin approach for reward systems where there are multiple measures of reward and contribution involved. They provide a numerical example which includes three reward measures (outputs) and two contribution measures (inputs). Despite the fact that there are only seven alternatives, maximin still does not produce a single winner. Looking at Fig. 1 would seem to indicate that in the case of two outputs there will normally be three alternatives which will appear at the extremes of the score range. This is because the frontier line needs to come in as close as possible to the data points in order to keep the score range narrow. In this case P gets the lowest score, with A and B getting the highest score. As the number of criteria are increased, the higher dimensionality of the problem means that the frontier will have more dimensions and so more observations will lie upon it. Hence, although having a single set of common criteria weights will reduce the number scoring 100%, we cannot rely on the maximin approach to produce a single winner.

4 Criteria Can be Completely Ignored

Consider the set of alternatives displayed in Fig. 2. In this case R will have the lowest score as it has the worst performance on both criteria. Its score will be maximised by referring to the horizontal dashed line as a frontier. R is not fully enveloped by a pair of frontier units in the way that P was in Fig. 1, and this causes difficulties. We shall now show that using the extension of this horizontal line as a frontier to assess all other alternatives leads to criterion 1 being completely ignored in the assessment i.e. a zero weight will be applied. The demonstration involves the similar right-angled triangles $R'Y_A O$, and $RY_R O$. The angle subtended at the origin is the same for both triangles, and the cosine of this angle equates to $OY_R/OR = OY_A/OR'$. Therefore $OR/OR' = OY_R/OY_A$. But OR/OR' is precisely the score for R and OY_R/OY_A is the ratio of values on criterion 2. Thus the values on criterion 1 play no part in the assessment of R. The same argument applies to the assessment of the other alternatives.

5 Conclusion

At first sight using the maximin rule to choose a set of common weights might seem an attractive approach to an analyst. One reason is that it is not subjective, but more importantly, it reduces the likelihood of being confronted by those who fare badly from the resulting rankings – this is because the method focuses on raising their score. Thus the analyst may be able to avoid having to argue with low scorers about the weights chosen.

However, this paper has shown that a number of serious drawbacks arise when using this rule in assessing performance. Any choice of weights corresponds to deciding how much each criterion is worth in terms of utility or value. It is clear that the maximin rule is allowing those who performed worst to effectively determine these utility values. This is as sensible as allowing the worst performing student to decide how much weight to attach to each of the various assessments taken by the class.

Next consider the problem of selecting from a set of alternatives. To ease the decision a common way to reduce the number of alternatives is to use screening or filtering. This is simply the removal of those alternatives which are clearly inadequate because they do not meet certain minimal standards. This step is carried out for convenience and should not affect the final decision. However, when used in conjunction with the maximin rule such a process will remove the worst performers and so lead to a different set of weights and a different ranking of the remaining alternatives. Decisions based on the maximin rule are highly sensitive to the inclusion or exclusion of alternatives whose performance is so poor as to be completely irrelevant to the selection decision.

We also showed that when the worst performing alternative is not naturally enveloped by units on the frontier (a common occurrence with real data), then certain criteria will be given zero weight and so be completely ignored in the analysis. Given that the criteria will have been carefully selected as being appropriate at the start, it is strange that they are now being dismissed.

Whilst, the maximin approach has been used in the allocation of resources in order to reduce inequality, its use to assess such a situation of need is a different matter entirely. The stage of evaluation to determine who is most in need or most deserving is separate from the stage of assigning resources or rewards. Rawls' difference principle may be of use in the allocation stage but not in the assessment stage. To persist in using it for both would be to minimise the apparent need of the worst off and thereby reduce the resources allocated to them.

References

- Butler M, Williams HP (2002) Fairness versus efficiency in charging for the use of common facilities. *J Opl Res Soc* 53:1324–1329
- Chernoff H (1954) Rational selection of decision functions. *Econometrica* 22:423–443
- Du Ding-Zhu (1996) Minimax and its applications. In: Horst R, Pardalos PM (eds) *Handbook of global optimization*. Kluwer, Dordrecht, pp 339–367
- Karsak EE (2004) A practical common weight MCDM approach for emerging market selection, *Proceedings of the 34th International Conference on Computers and Industrial Engineering (ICC&IE)*, San Francisco, California, USA, pp 159–164
- Karsak EE, Ahiska SS (2005) Practical common weight multi-criteria decision-making approach with an improved discriminating power for technology selection. *Int J Prod Res* 43(8):1537–1554
- Ogryczak W (1997) On the lexicographic minimax approach to location problems. *EJOR* 100: 566–585

- Pettypool MD, Karathanos P (2004) An equity check. *EJOR* 157:465–470
- Rardin RL (1998) *Optimization in operations research*. Prentice Hall, New Jersey
- Rawls JA (1999) *A theory of justice* (Revised edition). Harvard University Press, Cambridge, MA
- Sen AK (1969) Quasi-transitivity, rational choice and collective decisions. *Rev Econ Stud* 36:381–393
- Troutt MD (1997) Derivation of the maximin efficiency ratio model from the maximum decisional efficiency principle. *Ann Ops Res* 73:323–338
- Troutt MD, Zhang A (1993) The maximin efficiency ratio as a one step heuristic ranking device – comparison to two modified DEA analyses. Working paper, Dept of Management, Southern Illinois University, Carbondale, Illinois
- Troutt MD, Zhang A, Pettypool MD (1993) A maximin efficiency ratio model for consensus weights and extended ranking of technically efficient units. Working paper, Dept of Management, Southern Illinois University, Carbondale, Illinois
- Yaari ME, Bar-Hillel M (1984) On dividing justly. *Soc Choice Welfare* 1:1–24